Naval Research Laboratory

Washington, DC 20375-5000

NRL/FR/5620—92-9522

Dual Difference Filtering: A Replacement for Interpolation and Subtraction to Detect Changes in Misregistered Signals

ALAN P. SCHAUM

Advanced Concepts Branch
Optical Sciences Division

November 30, 1992

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave Blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	November 30, 1992	

4. TITLE AND SUBTITLE

Dual Difference Filtering: A Replacement for Interpolation and Subtraction to Detect Changes in Misregistered Signals

5. FUNDING NUMBERS

PE - 0602111N
PR - RL1A RA11W53 AB

6. AUTHOR(S)

Alan P. Schaum

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/FR/5620--92-9522

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of the Chief of Naval Research
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Dual Difference Filtering (DDF), a symmetric signal processor developed as a means of detecting changes in a pair of displaced digital signals, supplants interpolation and subtraction. For oversampled signals, DDF performance can be described by an error spectrum that is the product of the spectrum of the underlying continuous signal and an attenuation factor that depends only on the filters. For undersampled or stochastic signals, performance is similarly describable, but in an average sense.

Simple formulae are found for the attenuation factors in terms of the filters' real- or frequency-space descriptions. A fundamental equation for error minimization is also derived that is useful when the form of the power spectrum is known. Optimal DDFs are computed for a power-law spectrum that is appropriate for imagery. Also, flexible suboptimal design methods are developed that produce superior performance for a wide range of signal spectra. Typical signal-to-clutter improvements of 30 dB relative to standard interpolation/subtraction methods are demonstrated.

14. SUBJECT TERMS

Change detection
Interpolation
Digital Filtering
Linear estimation
Nyquist sampling
Moving target indication
Image processing

15. NUMBER OF PAGES

39

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102

CONTENTS

DUAL DIFFERENCE FILTERING: A REPLACEMENT FOR INTERPOLATION AND SUBTRACTION TO DETECT CHANGES IN MISREGISTERED SIGNALS

1. INTRODUCTION AND BACKGROUND

This report introduces an alternative to interpolation and subtraction as a means of detecting differences in a pair of digital signals. A common application for this method of change detection is to digitized imagery in, for example, remote sensing [1] or medical imaging [2,3] but it is also useful in time delay estimation and in many other traditional signal processing problems [4]. The algorithmic technique proposed here can relax optical/mechanical alignment tolerances for satellite-based systems, reducing a vital cost/risk factor [5]. Improvements in the quality of radiological images using these methods can allow "reduction of injected contrast agents or X-ray doses" [6]. The new methods also significantly enhance signal-to-clutter ratios in autonomous surveillance applications and in military search and track systems [7] in which the detection of moving targets is based on digital background subtraction followed by thresholding.

The standard procedure for detecting changes in these applications is to interpolate the first discrete signal at points corresponding to the samples of the second, and then subtract the signals. The shift s between the pair of sampling grids defining the discrete signals is assumed to be known deterministically or from signal-based algorithms [8]. The interpolator applied to the first signal usually takes the form of a convolution with a finite-extent kernel k. When evaluated at points shifted s from the original grid, the discrete result constitutes a resampled digital signal. The second signal is usually not resampled.

In an imaging application, Stocker and Oshagan [9] recognized that the spectral modification to the first signal induced by interpolation was lacking in the unaltered second signal. The resulting spectral mismatch limited the amount of clutter reduction in the difference signal. They used a polynomial resampler, the cubic B-spline (CBS), to partially overcome this problem. CBS has the unusual, noninterpolative property of not reproducing the original discrete signal at the sample points, i.e., in the limit $s \rightarrow 0$. Therefore, resampling the second signal with CBS at zero shift produces some spectral shaping. Subtraction from the first signal, resampled at shift s, then achieves a marked reduction in residual clutter.

One way to eliminate spectral mismatch completely is to apply a standard symmetric interpolator to the first signal at the shift $s/2$, and to the second at $-s/2$. However, better performance can be realized if the restrictive idea of interpolating is discarded. The goal of interpolation is to reproduce an underlying continuous signal. It is, therefore, appropriate when a comparison is made to an unaltered second signal. But if the actual goal is change detection, this approach is unnecessarily limiting.

This report considers the general effects of filtering both signals prior to subtraction and develops a mathematical method for evaluating and optimizing these Dual Difference

Manuscript approved September 14, 1992.

Filters (DDFs). Conventional interpolation is seen as a special case, when the second signal is not filtered. Interpolators with nontrivial behavior at zero shift like the cubic B-spline are seen as suboptimal types of DDF.

2. ANALYSIS

With $f(x)$ representing the underlying signal intensity at position x, sample values of $f, f(n)$ (where n is an integer) constitute the first digital signal, which may be convolved with a kernel $k_1^{(s)}$ to produce a discrete sequence

$$\sum_n k_1^{(s)}(m-n) f(n) . \tag{1}$$

In practical applications, this sum is finite because $k_1^{(s)}(m)$ is nonzero for only a few values of m near $m = 0$. The dependence of k_1 on s arises because the choice of filter generally depends on the separation of the sampling grids.

Similarly, the second digital signal $f(n + s)$, which is a sampled version of f at the shifted points, can be filtered to produce a second sequence, using a (possibly) different kernel $k_2^{(s)}$:

$$\sum_n k_2^{(s)}(m-n) f(n+s) . \tag{2}$$

It is convenient to redefine k_1 to allow for continuous arguments

$$k_1(n+s) \equiv k_1^{(s)}(n) , \tag{3}$$

where n is an integer. When Eq. (3) is substituted into expression (1), the filtered sequence in (1) assumes the form of a convolution evaluated at the point $m + s$, which is already the form in (2). The sequences in expressions (1) and (2) now can be treated as sample values of continuous functions f_1, f_2, where

$$f_1(x) \equiv \sum_n k_1(x-n) f(n) \tag{4a}$$

and

$$f_2^{(s)}(x) \equiv \sum_n k_2^{(s)}(x-n-s) f(n+s) , \tag{4b}$$

that are evaluated at $x = m + s$. Equation (4a) is a standard form for an interpolation function f_1 produced by a kernel k_1.

We will usually assume that $0 \le s \le 1/2$. That is, the second signal is shifted to the right of the first by no more than half a sample. This can always be arranged by relabelling the integer sample points and/or interchanging the order of the signals, if necessary. Also, the treatment here is in one dimension only. Higher-dimensional filtering can be accomplished by a cascade of two one-dimensional operations or, more generally, by treating the arguments x, n, and s as vectors.

In conventional change detection, before a subtraction is made, one of the signals—say, the first—is interpolated, and the other is unaltered. For this method, a good kernel k_1 produces an f_1 that is close to f, and the second kernel is trivial:

$$k_2^{(s)}(m-n) \rightarrow \delta_m^n , \tag{5}$$

where δ is the Kronecker delta. However, the requirement that k_1 be a good interpolator is too restrictive if the real goal is change detection. Nevertheless, a review of the interpolation application is instructive.

Because the sampling grids $\{n\}$ and $\{n+s\}$ are defined here as having unit spacing, the Nyquist frequency is $\nu_{Nyq} = 1/2$ cycles/sample. If the Fourier transform of f has no components with frequency $\nu \geq 1/2$, f is called "oversampled." According to the Nyquist Reconstruction Theorem, the kernel

$$k_1(x) = \text{sinc}(x) \equiv \frac{\sin(\pi x)}{\pi x} \tag{6}$$

may be used in Eq. (4a) to reproduce f exactly (i.e., $f_1 = f$), whenever f is oversampled.

In practice, k_1 must be of finite support (the range of x over which k_1 is nonzero) to make expression (1) or Eq. (4a) contain only a finite number of terms. For example, the sinc function can be truncated at $\pm N/2$ to make an N-point interpolator. Two other common interpolators are nearest neighbor (NN) and linear (LIN). Figure 1 illustrates $k_1(x)$ for NN, LIN, and SINC. The supports for NN and LIN are 1 and 2, respectively. The SINC has infinite support, but when truncated at $\pm N/2$, its support or *order* becomes N.

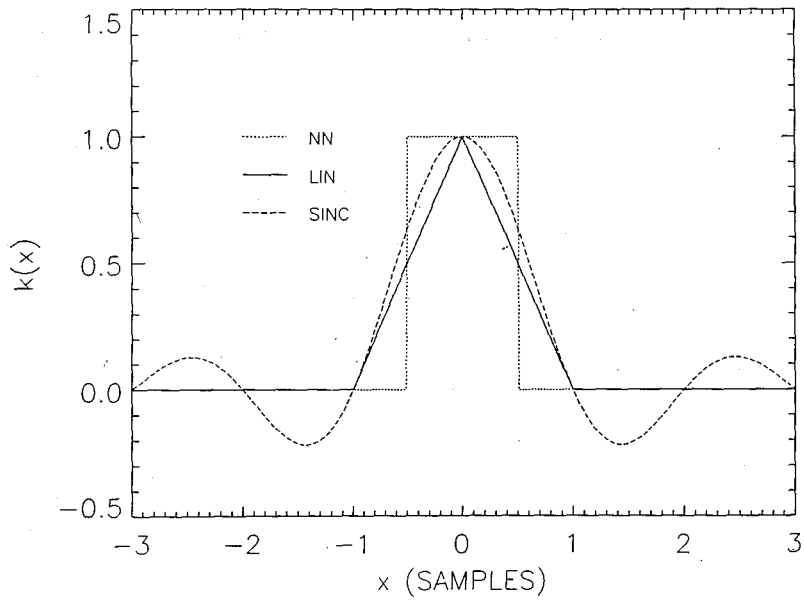

Fig. 1 — Kernels for three interpolators

The method of analysis developed here relies heavily on Fourier transform theory. The transform operator is denoted by $\mathcal{F}$, and upper and lower cases denote transform pairs. Also, the symbol $\Leftrightarrow$ indicates a correspondence between real and transform space: for example,

$$\left(\mathcal{F}k\right)(\nu) \equiv K(\nu) \equiv \int e^{-i\omega x}\, k(x)\, dx \qquad (\omega \equiv 2\pi\nu) \tag{7}$$

means the same thing as

$$k \Leftrightarrow K\,. \tag{8}$$

The shift and convolution theorems are then

$$\begin{aligned} g(x+s) &\Leftrightarrow e^{2\pi i s\nu}\, G(\nu) \\ e^{-2\pi i x\nu_o}\, g(x) &\Leftrightarrow G(\nu+\nu_o)\,, \end{aligned} \tag{9}$$

and

$$\begin{aligned} g(x)\, h(x) &\Leftrightarrow (G * H)(\nu) \\ (g * h)(x) &\Leftrightarrow G(\nu)\, H(\nu)\,. \end{aligned} \tag{10}$$

Another important theorem [10,11] relates the comb function, an infinite sum of Dirac delta functions,

$$\mathrm{comb}(x) \equiv \sum_n \delta(x-n)\,, \tag{11}$$

to its transform

$$\mathrm{comb}(x) \Leftrightarrow \mathrm{COMB}(\nu)\,. \tag{12a}$$

Equations (11), (12a), and (7) immediately imply that

$$\mathrm{COMB}(\nu) = \sum_n e^{-2\pi i n\nu}. \tag{12b}$$

Another important relation that follows directly from the definitions of the continuous convolution operation $*$ and the comb is

$$\sum_n g(x-n)\, h(n) = [g * (\mathrm{comb}\cdot h)](x)\,, \tag{13}$$

which is valid for arbitrary g, h. A special case of Eq. (13) is also useful:

$$\sum_n g(x-n) = (g * \mathrm{comb})(x)\,. \tag{14}$$

These relations can be used to express a discrete sum of squared sample values of a function g in terms of its Fourier transform G. Letting $g \to g^2$, $x \to 0$, and $n \to -n$ in Eq. (14) results in

$$\sum g^2(n) = \left(g^2 * \mathrm{comb}\right)(0)\,. \tag{15}$$

Next, successively using Eqs. (10) and (12a), the definition of $\mathcal{F}^{-1}$, and Eq. (10) again in Eq. (15) produces

$$\begin{aligned}\Sigma g^2(n) &= \left[\mathcal{F}^{-1}\left(\mathcal{F}\left(g^2 * \text{comb}\right)\right)\right](0) = \left[\mathcal{F}^{-1}\left(\mathcal{F}\left(g^2\right)\text{COMB}\right)\right](0) \\ &= \int d\nu' \left[\mathcal{F}\left(g^2\right)(\nu')\,\text{COMB}(\nu')\right] = \int d\nu' \left[G * G\right](\nu')\,\text{COMB}(\nu')\,. \end{aligned} \quad (16)$$

The frequency-space version of Eq. (11) is now used in Eq. (16) along with the definition of convolution:

$$\begin{aligned}\Sigma g^2(n) &= \sum_n \int d\nu'\, \delta(\nu' - n) \int d\nu\, G(\nu' - \nu)\; G(\nu) \\ &= \sum_n \int d\nu\, G(n - \nu)\; G(\nu)\,. \end{aligned} \quad (17)$$

We assume for convenience that g is real. Then

$$G(\nu) = G^*(-\nu) \qquad \forall\, \nu\,, \quad (18)$$

and from Eq. (17),

$$\Sigma g^2(n) = \sum_n \int d\nu\, G^*(\nu - n)\; G(\nu)\,. \quad (19)$$

If, furthermore, the sequence $g(n)$ results from oversampling the continuous function g, so that $G(\nu) = 0$ for $|\nu| \geq 1/2$, then Eq. (19) simplifies to

$$\Sigma g^2(n) = \int d\nu\, |G(\nu)|^2\,, \quad (20)$$

which is closely related to Parseval's Theorem, with the usual roles of position and frequency interchanged. Note the distinction, however: in Eq. (20), G is the Fourier transform of the underlying function g, not of the Fourier series $\{g(n)\}$. These two transforms are equal if g is oversampled, but even if it is not, Eq. (20), which we will call the discrete Parseval's Theorem, still holds in an average sense described below.

Note from Eq. (20) that for any oversampled g, $\Sigma g^2(n)$ is independent of where the sampling grid is laid down relative to the continuous signal $g(x)$. Shifting the grid by an amount t in one direction is equivalent to shifting $g(x)$ in the other; i.e.,

$$g(x) \rightarrow g(x + t) \quad (21)$$

which, according to Eq. (9), in Fourier space is equivalent to

$$G(\nu) \rightarrow G(\nu)\, e^{2\pi i t \nu}, \quad (22)$$

under which change the right side of Eq. (20) is invariant. Because the value of $\sum g^2(n)$ is independent of grid placement for an oversampled g, we call Eq. (20) the strong form of the discrete Parseval's Theorem. When g is not oversampled, a weaker version still holds, in which g is randomly shifted with respect to the sampling grid.

In particular, under the translation of the function g by the amount $t \in [0,1]$, according to Eq. (22), Eq. (19) becomes

$$\sum g^2(n) \rightarrow \sum_n e^{2\pi i t n} \int d\nu \; G^*(\nu - n) \; G(\nu) . \tag{23}$$

Averaging Eq. (23) over t: $0 \rightarrow 1$ yields Eq. (20). That is, the weak form of the discrete Parseval's Theorem is also expressed by Eq. (20), if the left-hand side is understood as an average over all placements of the sampling grid.

Methods similar to those above can also be used to generalize both forms of the discrete Parseval's Theorem and to derive a simple formula characterizing the performance of any DDF. To compute the differencing error associated with the DDF (k_1, k_2), we let

$$g^{(s)}(x) = f_1(x+s) - f_2^{(s)}(x+s) \tag{24}$$

(see Eq. (4)). Then the squared error in the difference signal is

$$d_s^2 \equiv \sum_n \left[f_1(n+s) - f_2^{(s)}(n+s) \right]^2 , \tag{25}$$

or

$$d_s^2 = \sum \left[g^{(s)}(n) \right]^2 . \tag{26}$$

Next, from Eqs. (24) and (9),

$$G^{(s)}(\nu) = \left[F_1(\nu) - F_2^{(s)}(\nu) \right] e^{2\pi i \nu s} . \tag{27}$$

The asymmetry in the (s)-dependence introduced in Eqs. (3) and (4) accounts for the corresponding asymmetry in the F_i of Eq. (27). This notation was introduced to facilitate use of the discrete Parseval's Theorem.

If the sampling grid were dense enough that $g^{(s)}(n)$ represented an oversampling of $g^{(s)}(x)$, then Eq. (20) could be applied immediately to Eqs. (26) and (27), with $g \rightarrow g^{(s)}$. But the functions f_i defining $g^{(s)}$ in Eq. (24) are usually not bandlimited, and so neither is $g^{(s)}$, because the transformations of Eq. (4) introduce aliasing whenever either kernel k_i is of finite support. Therefore, g cannot be oversampled, even if the f_i are, and the strong Parseval's Theorem does not apply.

Nevertheless, when Eqs. (24), (4a), and (4b) define g, Eq. (20) still holds whenever f is oversampled, even if g is not. This key property will permit the simple characterization of the error (Eq. (25)) as a power spectrum that is the product of an attenuation factor and the power spectrum of f. The attenuation factor depends only on the kernels k_i and the shift s.

Using Eq. (13) in Eq. (4a) with $h = f$ and $g = k_1$, taking the transform, and applying Eq. (10) results in

$$F_1(\nu) = K_1(\nu)[\mathrm{COMB} * F](\nu), \tag{28}$$

which is usually not bandlimited because neither factor is, as will be demonstrated. Consequently, G in Eq. (27) is usually not bandlimited.

The frequency version of Eq. (14) allows us to rewrite Eq. (28) as

$$F_1(\nu) = K_1(\nu) \sum_m F(\nu - m). \tag{29}$$

By similar means, it can be shown that

$$F_2^{(s)}(\nu) = K_2^{(s)}(\nu) \sum_m F(\nu - m)\, e^{-2\pi i m s}. \tag{30}$$

Even if F is bandlimited, the second factor in Eq. (29) or Eq. (30) is nonzero for arbitrarily large ν. So is the first whenever the kernel $k_i^{(s)}$ is of finite support, i.e., in all practical cases. Consequently, g is not bandlimited, as claimed, because the f_i are not.

Nevertheless, using Eq. (26) along with Eqs. (19), (27), (29), and (30) results in

$$d_s^2 = \sum_n \int d\nu \sum_m F^*(\nu - n - m)\left[K_1(\nu - n) - e^{-2\pi i s m} K_2^{(s)}(\nu - n)\right]^* \times$$
$$e^{2\pi i s n} \sum_j F(\nu - j)\left[K_1(\nu) - e^{-2\pi i s j} K_2^{(s)}(\nu)\right], \tag{31}$$

which we now show to be independent of grid placement whenever f is oversampled.

Assuming then that

$$F(\nu) = 0 \qquad \text{for} \quad |\nu| \geq \frac{1}{2}, \tag{32}$$

we examine a typical cross-term of Eq. (31):

$$-\sum_n e^{2\pi i s n} \int d\nu \,[K_1(\nu - n)]^* \sum_m F^*(\nu - n - m)\; K_2^{(s)}(\nu) \sum_j F(\nu - j)\, e^{-2\pi i s j}. \tag{33}$$

Because of Eq. (32), the integration in Eq. (33) kills all summand terms except when $n + m = j$. Notice also that averaging over grid placement t, as described in Eqs. (21) and (22), has the same effect. This means that the following results hold in the weak sense also, for which Eq. (32) need not apply.

Equation (33) thus reduces to

$$-\sum_n \int d\nu\,[K_1(\nu-n)]^* \sum_m |F(\nu-n-m)|^2\, K_2^{(s)}(\nu)\; e^{-2\pi i s m} \quad . \tag{34}$$

Changing variables, $\nu \rightarrow \nu + n + m$ in Eq. (34) produces

$$-\int d\nu\,|F(\nu)|^2 \left\{ \sum_{n,m} e^{-2\pi i s m} [K_1(\nu+m)]^*\, K_2^{(s)}(\nu+n+m) \right\} . \tag{35}$$

The other cross-terms of Eq. (31) produce results similar to expression (35), but with different terms within the braces. The final result may be written as

$$d_s^2 = \int |F(\nu)|^2 |E_s(\nu)|^2\, d\nu \;, \tag{36}$$

which includes an attenuation factor that is the squared magnitude of a "complex error factor" $E_s(\nu)$:

$$E_s(\nu) \equiv \sum_n \left[e^{2\pi i n s} K_1(\nu+n) - K_2^{(s)}(\nu+n) \right] . \tag{37}$$

The quantity $|F(\nu)|^2|E_s(\nu)|^2$ will be called the error spectrum. The error factor $E_s(\nu)$ depends only on the kernels, and not on the signals being processed. The independence of Eq. (36) from the phase of F means that d_s^2 does not depend on the placement of the sampling grids when the signal is oversampled, even though the individual terms of Eq. (25) do. Furthermore, as indicated earlier, Eq. (36) holds also for undersampled signals when averaged over all such placements. Therefore, the result expressed by Eq. (36) has strong and weak forms, just as the discrete Parseval's Theorem has (Eq. (20)).

The formula in Eq. (37) is most useful when the sum contains only a finite number of terms, that is when the Fourier transforms of the kernels are of finite support. For example, for perfect (inband) interpolation $k_1(x) = \mathrm{sinc}(x)$, and $K_1(\nu)$ has the value 1 inband ($-1/2 \le \nu \le 1/2$), and is zero out of band. If the second signal is left unaltered, then $k_2^{(s)}(x)$ can be chosen to be any function satisfying Eq. (5); a convenient choice is $k_2^{(s)}(x) = \mathrm{sinc}(x)$. Then $K_2(\nu) = K_1(\nu)$, and Eq. (37) can be used to show that for sinc interpolation the error factor is given by

$$|E_s(\nu)| = 2\left| \sin\left(\pi s\, \mathrm{int}\left[\nu + \frac{1}{2}\right]\right)\right| , \tag{38}$$

where "int" means "largest integer less than or equal to." (For example, int[−.5] = −1.) Thus, the error factor is zero inband, but not out of band, where many other interpolators are superior [12].

In many practical situations, it is the real-space form of a kernel that is of finite support, and so the next goal is to express Eq. (37) in terms of k_i instead of K_i.

The first term of Eq. (37) is, by definition,

$$\sum_n e^{2\pi i n s} \int k_1(x)\, e^{-2\pi i x(\nu+n)} dx\,. \tag{39}$$

The real-space version of Eq. (12b) along with Eq. (11) then permits us to rewrite Eq. (39) as

$$\int k_1(x)\, e^{-2\pi i x\nu}\, \mathrm{comb}(s-x)\, dx = \sum_n e^{-2\pi i \nu(s+n)}\, k_1(s+n)\,. \tag{40}$$

A similar expression for the second term of Eq. (37) leads to

$$E_s(\nu) = \sum_n e^{-i\omega n}\left[e^{-i\omega s}\, k_1(s+n) - k_2^{(s)}(n)\right] = \sum_n e^{-i\omega n}\left[e^{-i\omega s}\, k_1^{(s)}(n) - k_2^{(s)}(n)\right], \tag{41}$$

where Eq. (3) has also been used. Equation (41) is an alternate form for Eq. (37); it expresses the error factor directly in terms of the kernels instead of their Fourier transforms.

According to expressions (1) and (2), the values $k_1^{(s)}(n)$ and $k_2^{(s)}(n)$ are tap weights of digital filters operating on the pair of discrete signals $f(n)$ and $f(n+s)$. Calling the transfer functions of these filters $H_1^{(s)}$ and $H_2^{(s)}$, i.e.,

$$H_1^{(s)}(\nu) \equiv \sum e^{-i\omega n}\, k_1^{(s)}(n)\,, \tag{42a}$$

$$H_2^{(s)}(\nu) \equiv \sum e^{-i\omega n}\, k_2^{(s)}(n)\,, \tag{42b}$$

the complex error factor may be written compactly as

$$E_s(\nu) = e^{-i\omega s}\, H_1^{(s)}(\nu) - H_2^{(s)}(\nu)\,. \tag{43}$$

Note that the sums in Eqs. (42a) and (42b) are finite in practical applications. In the conventional method of change detection, in which the first signal is interpolated and the second is unaltered, Eq. (5) applies, and the general error factor for interpolation/subtraction reduces to

$$E_s(\nu) = e^{-i\omega s}\, H_1^{(s)}(\nu) - 1\,. \tag{44}$$

The form for interpolation error that results from inserting Eq. (44) into Eq. (36) is equivalent to one derived in Ref. 13 for the special case of a purely sinusoidal input signal with a sub-Nyquist frequency.

As an example, it can be shown that the perfect inband interpolator (Eq. (6)) has transfer function

$$H_1^{(s)}(\nu) \rightarrow e^{2\pi i s\,(\nu - \mathrm{int}[\nu + \frac{1}{2}])} \quad \text{(sinc interpolation)}. \tag{45}$$

Equation (38), which was derived from the continuous-transform representation in Eq. (37), can be recovered by substituting Eq. (45) into Eq. (44).

Equations (36), (37), and (41) or (43) express the resampling error at any fixed shift s between a pair of sampling grids in terms of the spectrum of the underlying continuous signal and an error factor that depends only on the method of filtering in expressions (1) and (2). Furthermore, if f is not oversampled, then Eq. (36), which constitutes a fundamental theorem of resampling error, still holds in the weak sense, i.e., averaged over all possible placements of the two sampling grids relative to the signal. Finally, if the function f is regarded as one realization of a stationary stochastic process, then taking the ensemble average of Eq. (36) yields a third form of the theorem, which relates the mean squared error to $|E|^2$ and $P(\nu) = \langle |F(\nu)|^2 \rangle_{\text{ensemble}}$, which is the power spectrum of the stochastic process.

3. OPTIMAL DUAL DIFFERENCE FILTERS

The three forms of the fundamental theorem can be used to design DDFs satisfying a variety of criteria, such as error minimization. For example, Eqs. (36) and (41) may be combined to give

$$d_s^2 = \sum_{m,n} R(n)\left[k_1(n+m+s)\,k_1(m+s) + k_2^{(s)}(n+m)\,k_2^{(s)}(m)\right] - 2\sum_{m,n} R(n+s)\,k_1(n+m+s)\,k_2^{(s)}(m) \tag{46}$$

with

$$R(x) \equiv \int_{-\infty}^{\infty} \cos(x\omega)\,|F(\nu)|^2\,d\nu. \tag{47}$$

For the interpolation application (see Eq. (5)), this reduces to

$$d_s^2 = R(0) - 2\sum_n R(n+s)\,k_1(n+s) + \sum_{m,n} R(n)\,k_1(n+m+s)\,k_1(m+s). \tag{48}$$

Minimization of d_s^2 in Eq. (48) with respect to the variables $k_1(n+s)$ results in an equation for the optimal N-point interpolation kernel

$$\sum_m R(m)\,k_1(n-m+s) = R(n+s), \tag{49}$$

valid for those n for which $k_1(n+s)$ in Eq. (48) is nonzero, i.e., for the N values of n: int$[-(N-1)/2]$,..., int$[(N-1)/2]$. Comparing Eq. (49) with Eq. (4a) shows that the former is equivalent to the requirement that k_1 serve as a perfect interpolator for the function $R(x)$ for $x \in [-N/2, N/2]$, i.e., over the support defining the order of k_1. Equation (49) has appeared before [1], but only as the condition minimizing an *ensemble*

mean-squared error for a stationary stochastic process, that is, in a more limited application than shown here.

Because $R(m)$ is symmetric, Eq. (49) is a Töplitz equation, for which general iterative solution procedures have been developed [14,15]. Thus, in principle, $k_1(s+n)$ can always be found as a function of $2N$ numbers, $R(m)$ and $R(n+s)$, with $m = 0,..., N-1$ and $n = \text{int}[-(N-1)/2]$,..., $\text{int}[<(N-1)/2]$. Solutions to Eq. (49) were found in Ref. 12 for various power spectral models: uniform, Gaussian, power law, and Lorentzian.

Error minimization equations can be derived for the more general case of DDFs in Eq. (46), but now at least one extra constraint must be added, lest the trivial solution $k_1(n+s) = k_2^{(s)}(n) = 0$ result. The nature of such constraints depends on that of the change that is to be detected. It may be an extensive change, such as in a region of blood flow on an angiogram, or large-scale agricultural evolution as seen from Earth orbit. It may be a local change, such as the motion of a small object. The general goal is to combine the two signals after filtering so that only such changes persist.

In dual difference filtering, at least one of the individual filters k_i must be designed to preserve the change that is to be detected. As an example, for extensive changes, an appropriate choice of constraint is to require perfect dc response of each filter. This will tend to preserve local averages. Similarly, for detection of small moving targets of known shape, the filters may be constrained to conserve target energy or peak amplitude.

Generally, the inclusion of constraints in the DDF optimization problem can be achieved by introducing Lagrangian multipliers λ_i into Eq. (46):

$$d_s^2 \to \sum_{m,n} R(n)\left[k_1(n+m+s)\,k_1(m+s) + k_2^{(s)}(n+m)\,k_2^{(s)}(m)\right]$$
$$-2\sum_{m,n} R(n+s)\,k_1(n+m+s)\,k_2^{(s)}(m) - 2\sum_i \lambda_i c_i\left(\left\{k_1(n+s),\, k_2^{(s)}(n)\right\}\right). \tag{50}$$

The functions c_i represent constraints on the digital filters. The error minimization equations then result from differentiating Eq. (50) with respect to the tap weights $k_1(n+s)$ and $k_2^{(s)}(n)$, and λ_i :

$$\sum_m R(m)\,k_1(n-m+s) - \sum_m R(m+s)\,k_2^{(s)}(n-m) = \sum_i \lambda_i \frac{\partial c_i}{\partial k_1(n+s)} \tag{51a}$$

$$\sum_m R(m)\,k_2^{(s)}(n-m) - \sum_m R(m+s)\,k_1(n+m+s) = \sum_i \lambda_i \frac{\partial c_i}{\partial k_2^{(s)}(n)}, \tag{51b}$$

and

$$c_i\left(\left\{k_1(n+s),\, k_2^{(s)}(n)\right\}\right) = 0\,. \tag{51c}$$

In many optimization problems, the solutions for k_1, $k_2^{(s)}$ show a symmetry reflecting the fact that neither of the two original discrete signals is preferred. The symmetry is

$$k_1(n+s) = k_2^{(s)}(-n)\,, \tag{52a}$$

or, from Eq. (3),

$$k_1^{(s)}(n) = k_2^{(s)}(-n)\,, \tag{52b}$$

which is illustrated in Fig. 2 for two- and three-point filters. In the remainder of this report, the symmetry in Eq. (52) will be assumed for all DDFs, and the subscript on k_1 will be omitted. Other symmetries are considered in Ref. 16.

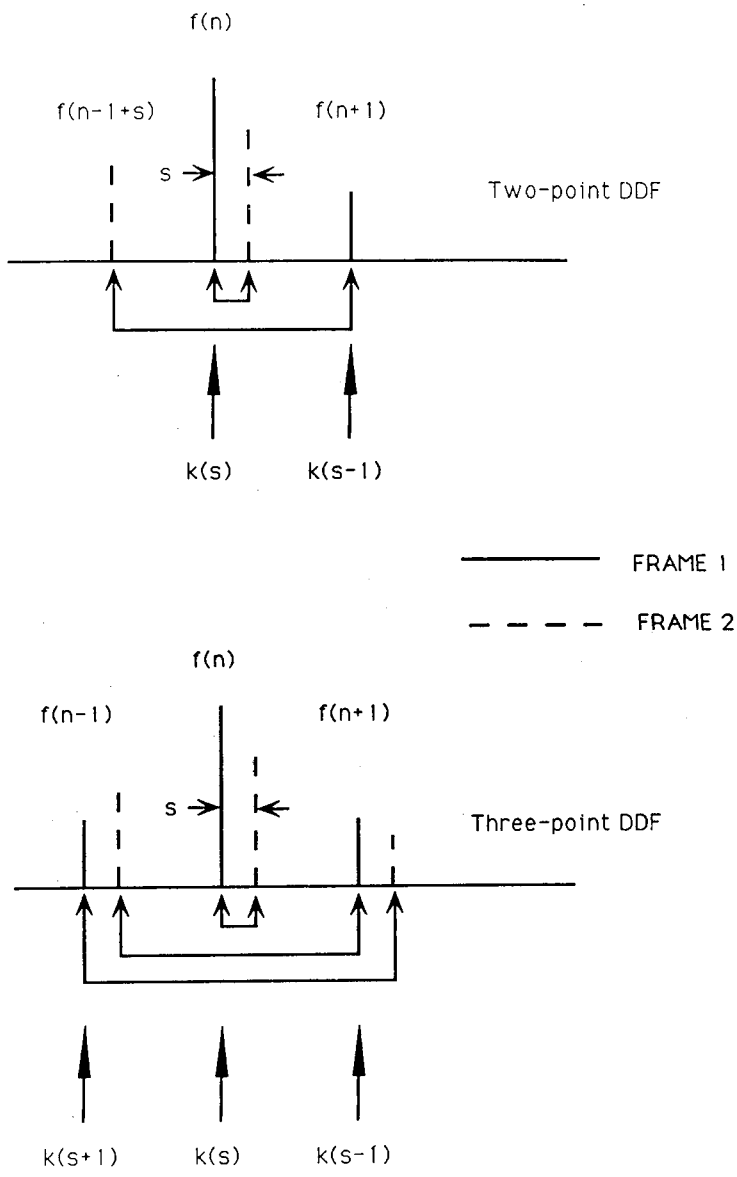

Fig. 2 — The usual symmetry in DDF filters. The double-arrow lines connect those sample points in the two digital signals that are multiplied by equal-value tap weights from the corresponding kernels.

For the symmetric case, the DDF error factor (Eq. (41)) reduces to

$$E_s(\nu) \rightarrow 2\sum_n k(s+n)\sin\left[\omega\left(n+\frac{s}{2}\right)\right], \tag{53}$$

in which an irrelevant unimodular phase factor has been omitted. Also, the residual error after subtraction (Eq. (46)) simplifies to

$$d_s^2 = 2\sum_{m,n}\{R(n)k(n+m+s)k(m+s) - R(n+s)k(n+m+s)k(-m+s)\}, \tag{54}$$

and now there are fewer independent optimization equations:

$$\sum_m R(m)k(n-m+s) - \sum_m R(m+s)k(m-n+s) = \sum_i \lambda_i \frac{\partial c_i}{\partial k(n+s)} \tag{55a}$$

$$c_i(\{k(n+s)\}) = 0 \tag{55b}$$

than in the general case (Eq. (51)).

The only type of constraint considered here in detail is that of a single linear one, which can be written as

$$c(\{k(n+s)\}) \equiv \sum_n k(n+s)t_{-n} - 1 = 0. \tag{56}$$

If $t_{-n} = 1$, Eq. (56) describes the "dc" constraint. Then the transfer functions in Eq. (42) have unit response at $\nu = 0$. Alternatively, $\{t_n\}$ may describe the amplitude of a target image whose interframe motion is to be detected by use of a DDF. Gaussian targets will be considered, with a standard deviation of σ samples and unit peak value

$$t_n = e^{-n^2/2\sigma^2}. \tag{57}$$

In all cases considered, t_0 has been normalized to unity. The values $\sigma = 1/2, 1$ will be used as examples. In imaging applications with matched optics design, these correspond respectively to undersampled and nearly Nyquist-sampled images of point objects [12].

The constraint in Eq. (56) then means that the peak signal value is reproduced by the filter (if the peak target value lies on the first sampling grid). It often happens that an asymmetrical form of the tap weights $\{k^s(n)\}$ transforms a symmetrical target into an asymmetrical one, with a consequent shift in the peak location. The filtered peak can then be larger than the unfiltered one. In any case, Eq. (56) requires that the peak value in the filtered signal be no less than in the unfiltered.

In all DDF examples studied here, error minimization requirements are satisfied simultaneously with a constraint that peak signal be conserved. These are always equivalent to a signal-to-clutter criterion. For example, requiring that the peak signal be doubled instead of conserved generally also doubles the filter tap weights and, from Eq. (46) or (54), the rms clutter $\sqrt{d_s^2}$ as well.

With the constraint in Eq. (56), Eq. (55a) becomes

$$\sum_m R(m)\,k(n-m+s) - \sum_m R(m+s)\,k(m-n+s) = \lambda\, t_{-n}\,, \tag{58}$$

valid for $n = \text{int}[-(N-1)/2]$,..., $\text{int}[(N-1)/2]$. These linear equations in N variables $k(n+s)$ should be compared to Eq. (49), the corresponding optimization equations for the interpolation problem. The explicit constraints in Eq. (56) are replaced there by a more severe one—the second signal is left unfiltered.

Before applying these results to specific problems, we note that as long as the quantities $R(m)$ and $R(m+s)$ are known, Eq. (58) can be used to find the optimal $k(n+s)$ even if the value of s is unknown, because all the s-dependence in Eq. (58) is implicit, appearing in the arguments of functions. This means that change detection might be possible without knowledge of the shift s between sampling grids, if estimates of $R(m)$ and $R(m+s)$ are available. And these may be provided by interpreting R as an autocorrelation function and then constructing periodograms [17]: (a) from either signal, to estimate $R(m)$, and (b) from both to estimate $R(m+s)$. This technique introduces errors by using estimates of R, but it completely eliminates errors from a prior step that we have ignored: estimation of the shift s. The competition between these errors depends on the performance of registration methods and is not studied here.

3.1 Power-Law Spectrum

A common analytic model for image power spectra is a low inverse power p of frequency, typically with $p \approx 2$. We will examine optimal DDFs for the spectrum

$$|F(\nu)|^2 \sim 1/\nu^2\,. \tag{59}$$

Besides its relevance to imagery, this choice is motivated by the simplicity of the optimal interpolator solution, which is known to all orders and can therefore be compared easily with the DDF solutions. Also, the method used here to handle the dc singularity is generalizable to other singular spectra. In practice, at some cutoff the low-frequency form of Eq. (59) becomes bounded, keeping the real-space image variance finite. The following analysis shows that the relative performance of most resamplers is insensitive to the cutoff.

We will treat Eq. (59) as the limiting case of the spectrum

$$|F(\nu)|^2 \sim \frac{1}{\nu^2+\varepsilon^2} \tag{60}$$

as $\varepsilon \to 0$. The function R in Eq. (47) is then easily evaluated:

$$R(x) \to \frac{\pi}{\varepsilon}\,\rho^{|x|}$$

where

$$\rho \equiv e^{-2\pi\varepsilon} \qquad (\varepsilon \geq 0)\,. \tag{61}$$

For two-point filtering, Eq. (58) reduces to

$$\begin{bmatrix} R(0)-R(s) & R(1)-R(1-s) & -1 \\ R(1)-R(1-s) & R(0)-R(2-s) & -t_1 \\ 1 & t_1 & 0 \end{bmatrix} \begin{bmatrix} k(s) \\ k(s-1) \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad (62)$$

whose solution is

$$\begin{aligned} k(s) &= m_{31} / det \\ k(s-1) &= -m_{32} / det \end{aligned} \quad (63)$$

with m_{ij} the cofactors of the matrix in Eq. (62) and *det* the determinant. In particular,

$$\begin{aligned} m_{31} &= R(0)-R(2-s)-t_1[R(1)-R(1-s)] \\ m_{32} &= R(1)-R(1-s)-t_1[R(0)-R(s)] \\ det &= m_{31} - t_1 m_{32} . \end{aligned} \quad (64)$$

The limit $\varepsilon \to 0$ corresponds to $\rho \to 1$, in which case $(\varepsilon/\pi)R(x) \to 1 \; \forall \; x$; Eq. (64) then shows that Eq. (63) is indeterminate. Application of l'Hôpital's Rule yields

$$\begin{aligned} k(s) &= [s(1-t_1)-2] / det \\ k(s-1) &= -s(1+t_1) / det \\ det &= 2(s-1)-s(1+t_1)^2, \end{aligned} \quad (65)$$

which describes the optimal two-point DDF for a $p = 2$ power-law spectrum. Except when t_1 assumes one of the two values $-1 \pm \sqrt{2}$, this kernel is a rational function of s, rather than a polynomial. Note that the optimal interpolation kernels for power-law spectra are polynomials [12], as are most interpolators in popular use. Thus, the DDF approach introduces a new class of kernels as a tool for change detection.

For the power spectrum of Eq. (59), the minimum-error interpolator, i.e., the solution to Eq. (49), is LIN for all $N \geq 2$ [1]. This is a peculiarity of the $p = 2$ spectrum, that the optimal interpolator has support 2 instead of N. Figure 3 plots the error factors $|E_s(\nu)|$ for each of three interpolators: the optimal (LIN), Cubic Convolution [18], which is a popular four-point interpolator, and the Low-Frequency optimal four-point interpolator (LF-4), which is designed to have the smallest possible error factor in the vicinity of $\nu = 0$ [12] and is, therefore, appropriate for the singular spectrum in Eq. (59). At $s = 0$, the errors are zero; at $s = 1/2$, they are usually the worst. Our examples are always at the intermediate shift $s = 1/4$, which choice explains (see Eq. (41)) the $\Delta\nu = 4$ periodicities in the error factors.

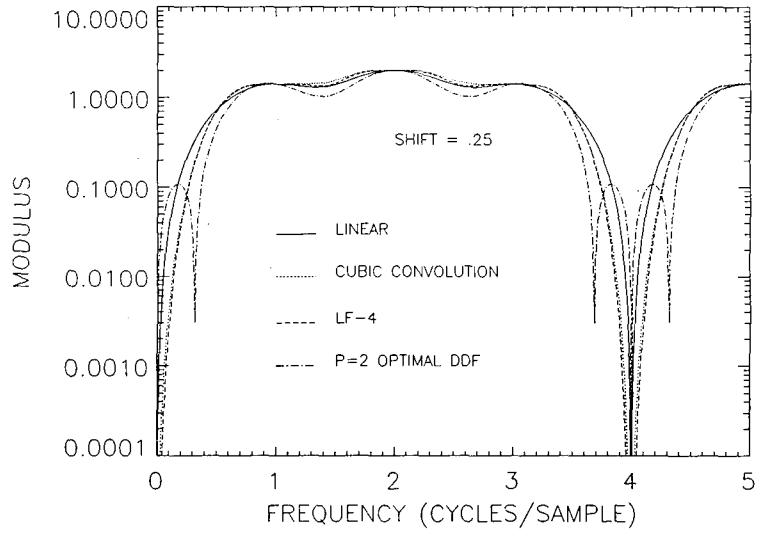

Fig. 3 — Error factors $|E_s(\nu)|$ at the shift $s = 0.25$ for three interpolators: Linear, Cubic Convolution, and LF-4; and one two-point DDF: the optimal for a $p = 2$ power law with the dc constraint. The Nyquist frequency is 0.5 cycles/sample

Figure 3 also shows the error factor for the optimal two-point DDFs in Eq. (65), using the dc constraint ($t_n = 1$ in Eq. (56)). Figure 4 shows the corresponding error spectra (the integrand of Eq. (36)) for the case of a $p = 2$ signal spectrum, and it lists the relative areas under these curves. By Eq. (36), these are also the relative mean squared errors in the filtered difference signals. Note that the optimal DDF, even for $N = 2$, outperforms all interpolators regardless of their order, because it outperforms the optimal, LIN. The polynomial interpolators (Cubic Convolution and LF-4) are included to illustrate their superior performance below the Nyquist frequency ($\nu = 1/2$), an important point in our later discussions.

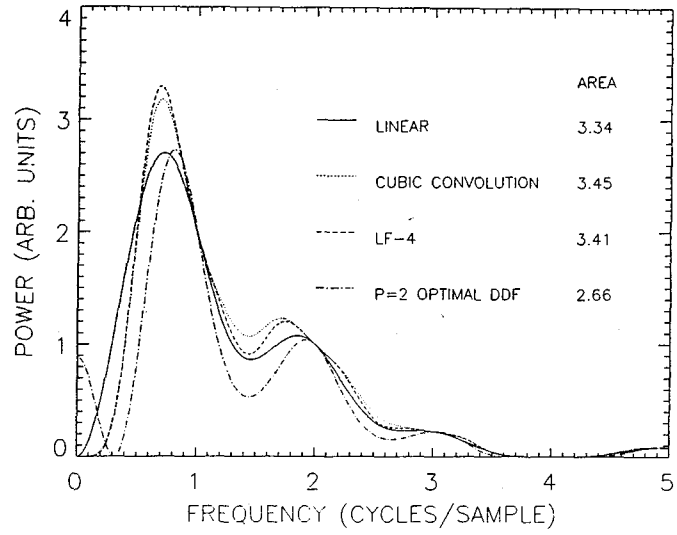

Fig. 4 — Error spectra $|E_s(\nu)\, F(\nu)|^2$ when the image spectrum is a $p = 2$ power law, for three interpolators and one DDF at the shift $s = 0.25$

Figures 5 and 6 show the effects of imposing on individual DDF filter performance the three types of constraint discussed earlier. Figures 7 and 8 show the corresponding results for the $p = 2$ optimal kernels when $N = 4$. As always, a shift $s = 1/4$ is assumed. The increasing errors with decreasing σ occur because smaller values of the standard deviation correspond to targets with a larger high-frequency content. Therefore, conserving the peak signal value competes increasingly with minimizing difference clutter, for which low frequencies are important for power-law spectra. The dc constraint may be considered a limiting version of the Gaussians as $\sigma \rightarrow \infty$.

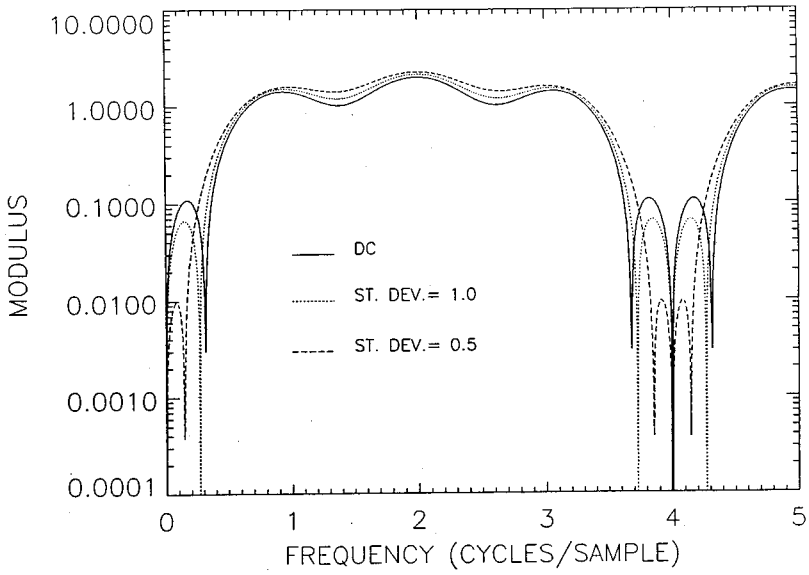

Fig. 5 — Error factors of the optimal two-point DDFs for a $p = 2$ power law, for three different filter constraints

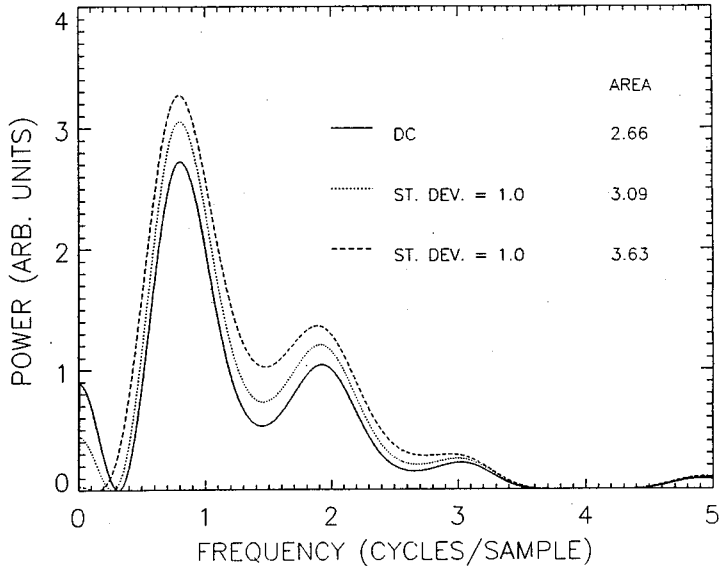

Fig. 6 — Error spectra after use of the optimal two-point DDFs, for three different filter constraints. The input signal spectrum is a $p = 2$ power law.

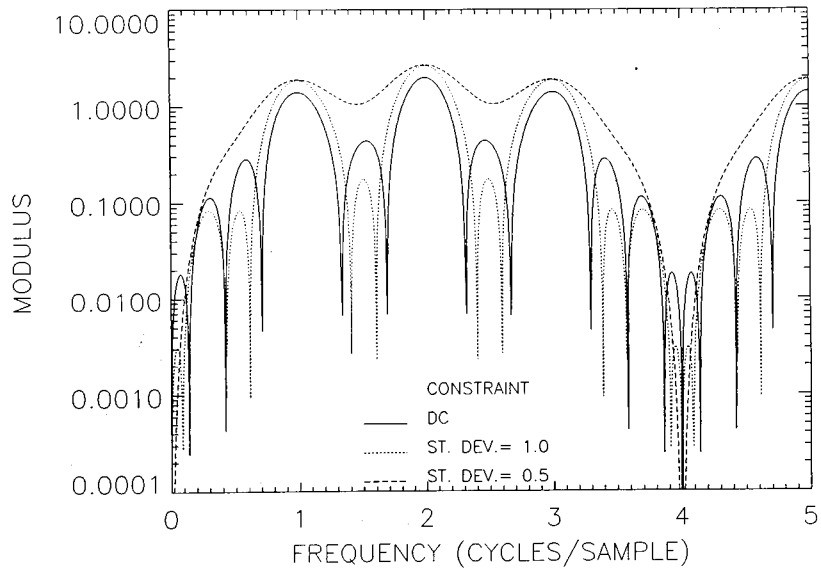

Fig. 7 — Error factors of the optimal four-point DDFs for a $p = 2$ power law, for three different filter constraints

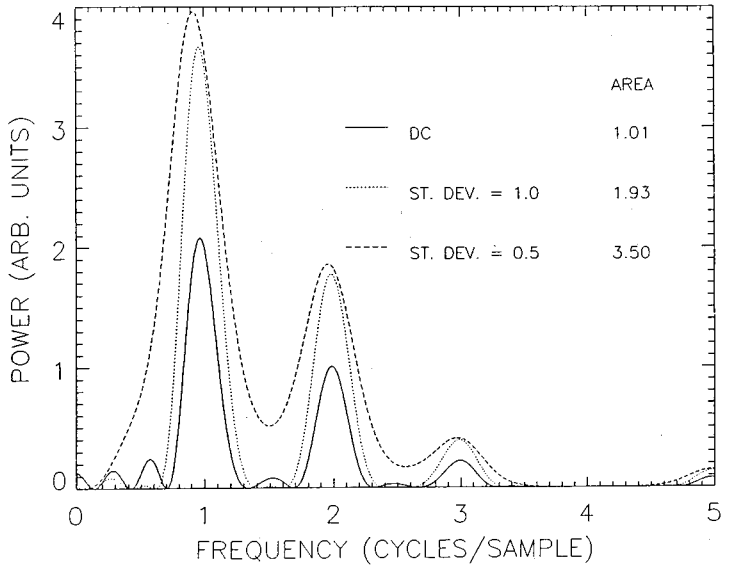

Fig. 8 — Error spectra after use of the optimal four-point DDFs, for three different filter constraints. The input signal spectrum is a $p = 2$ power law.

Notice that the frequency range plotted in these figures extends to 10 times the Nyquist frequency. If the power spectrum model of Eq. (59) is truncated to zero above the Nyquist frequency, then, as Figs. 3 and 4 show, the polynomial solutions are superior to the "optimal." Of course, the truncation produces a new $R(x)$, and a corresponding new optimal filter is described by Eq. (58). It can be found by a straightforward numerical procedure [16], which will not be pursued here. Instead, another, more robust approach that achieves near-optimal performance is described.

4. SUBOPTIMAL KERNEL DESIGN

Seldom does a simple analytic form like Eq. (59) or (60) describe all power spectra of interest. For imagery, for example, different power laws are often fit to the same spectrum in different frequency ranges. Also, a hard cutoff at some high frequency is generally imposed by the optical system collecting the data. A further complication is that a spectrum may differ significantly in different subregions of a single image. We now discuss methods for designing kernels that rely on more general properties of power spectra than a precise and simple analytic form.

4.1 Low-Frequency Optimal

The first example is Symmetric Low-Frequency optimal kernels of order N, called SLF-N. The performance of these DDFs will be compared to similar kernels, LF-N, found for the interpolation approach. Both types of kernel are defined by the requirement that the corresponding error factor $E_s(\nu)$ and the maximum possible number of its derivatives $E_s^{(m)}(\nu)$ be zero at $\nu = 0$. This criterion ensures a reasonably good performance for power spectra with nearly singular low-frequency behavior, which is common for imagery. The LF-N were studied in Ref. 12; their use is essentially a local version of the Lagrange interpolation procedure, which finds the unique $(N-1)$-order polynomial connecting N sample points.

For symmetric DDFs, Eq. (53) implies that

$$E_s^{(m)}(0) = 0 \tag{66}$$

for all even values of m. Therefore, if N_c is the number of constraints used to preserve the changes to be detected, then an N-point kernel has $(N - N_c)$ degrees of freedom, which can be used to satisfy Eq. (66) for several low odd values of m. Then Eq. (66) holds for $m = 0, \ldots, 2(N - N_c)$. (In our examples, N_c always has the value 1.)

By contrast, for interpolation, the N degrees of freedom can be used to make the error in Eq. (44) satisfy Eq. (66) for all s only for $m = 0, \ldots, N-1$. Therefore, at low frequencies, the error factor (Eq. (44)) associated with LF-N varies as ν^N, while for SLF-N with one constraint, it (Eq. (43)) varies as ν^{2N-1}. For example, for $N = 2$, the LF solution is just LIN, for which the error factor varies as ν^2; for SLF-2, derived below, it varies as ν^3. Thus, SLF-N produces less error at low frequencies than LF-N. This superior performance is analogous to the situation found above for the $p = 2$ optimal DDF, although there the DDF solution was even more preferred, in that even the $N = 2$ DDF kernel outperformed the optimal interpolator for any N.

To find the kernel for SLF-N with one constraint of the form in Eq. (56), we solve Eq. (66) which, from Eq. (53), becomes

$$\sum_n \left(n + \frac{s}{2}\right)^m k(n+s) = 0 \tag{67}$$

for $m = 1, 3, \ldots, 2N-3$. For example, for $N = 2$, the homogeneous Eq. (67) becomes

$$\left(-1 + \frac{s}{2}\right) k(-1+s) + \frac{s}{2} k(s) = 0 . \tag{68}$$

The inhomogeneous equation of constraint can be taken to be the dc condition

$$\sum_n k(n+s) = 1, \tag{69}$$

because the solutions $k(n + s)$ for the more general case of Eq. (56) are simply proportional to these. The same is true of the corresponding error factors $E_s(\nu)$, which are linear in $k(n + s)$. The dc solution can thus be used to express the general form of the peak-signal to rms clutter

$$S/C = \sum_n k(n+s)\, t_{-n} / d_s\,. \tag{70}$$

That is, if $k(n + s)$ is the SLF solution with the dc constraint (Eq. (69)), then Eq. (70) gives the S/C ratio for the more general problem defined by Eq. (56). The denominator is computed using Eqs. (36) and (41). Note that the analogous property does not generally hold for the optimal kernels for a given power spectrum. Figures 5 and 7 illustrate this, in that the ratio of the error factors corresponding to different constraints clearly depends on frequency, whereas Eq. (53) predicts the opposite for proportional kernels.

The solution to Eqs. (68) and (69) is

$$\begin{aligned} k(s) &= -\tfrac{1}{2}(s-2) \\ k(-1+s) &= \tfrac{1}{2}(s)\,. \end{aligned} \tag{71}$$

This has a simple interpretation in terms of conventional interpolation. Equation (71) represents the kernel for linear interpolation at a shift of $s/2$. From Eq. (52) then, the filtering of signal 2 consists of linear interpolation at a shift of $-s/2$. In other words, the SLF-2 solution amounts to shifting the two discrete signals toward each other, using linear interpolation, by half their separation. Such a simple relation does not, however, persist at higher orders for SLF-N.

The low-frequency optimal dc kernel solutions for SLF-3 are straightforward but tedious to derive. They satisfy two homogeneous equations, Eq. (67) with $m = 1, 3$, along with Eq. (69). The dc solution is

$$\begin{aligned} k(s-1) &= \tfrac{1}{12}(s+1)(s+2) \\ k(s) &= -\tfrac{1}{6}(s-2)(s+2) \\ k(s+1) &= \tfrac{1}{12}(s-1)(s-2)\,. \end{aligned} \tag{72}$$

For $N = 4$, the dc solution is

$$k(s-2) = \frac{1}{120}(s)(s+1)(s+2)$$

$$k(s-1) = -\frac{3}{120}(s-4)(s+1)(s+2)$$

$$k(s) = \frac{3}{120}(s-4)(s-3)(s+2)$$

$$k(s+1) = -\frac{1}{120}(s-2)(s-3)(s-4)\,. \tag{73}$$

The pattern of the monomial factors appearing in Eqs. (71) through (73) makes the use of an ansatz a feasible approach for higher N. For example, when $N = 5$, we expect fourth-order polynomial kernels $k(s-n)$ with monomial factors of the form $(s-m)$ with integer $|m| \leq 4$, $m \neq n$. The coefficients multiplying the various polynomials can be determined from the condition in Eq. (69).

Figure 9 includes the error factors for the above kernels along with those for the interpolators LF-2 (LIN) and LF-4. Note that the three-point DDF is better than the four-point interpolator at all inband frequencies.

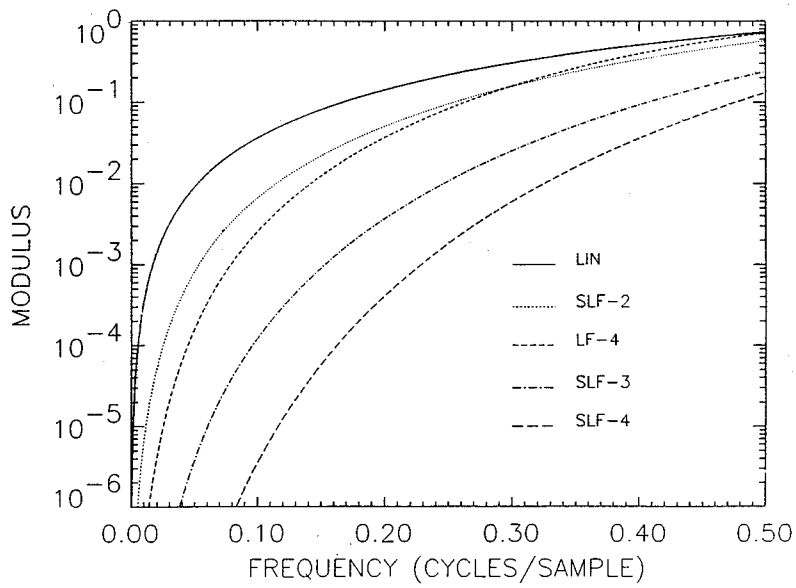

Fig. 9 — Error factors for two interpolators: Linear and LF-4; and three DDFs: SLF-2, 3, and 4

Figures 10 and 11 show the SLF-2, 3, and 4 error spectra for $p = 2$, 4 power laws, respectively. Note that no meaningful comparison can be made between Figs. 10 and 11 because no low-frequency cutoffs have been defined that could make the input spectra have equal areas and, hence, equal (and finite) unfiltered variances. For $p = 2$, the rms difference d_s is approximately 3.3 times smaller for SLF-3 than for SLF-2; and 7.1 times smaller for SLF-4 than for SLF-2. For $p = 4$, they are 3.8 and 8.7. (For $p = 3$, the corresponding ratios are 3.5 and 7.7, respectively.) The more singular spectra benefit relatively more as N increases for the SLF filters, which are designed to work well at low frequencies.

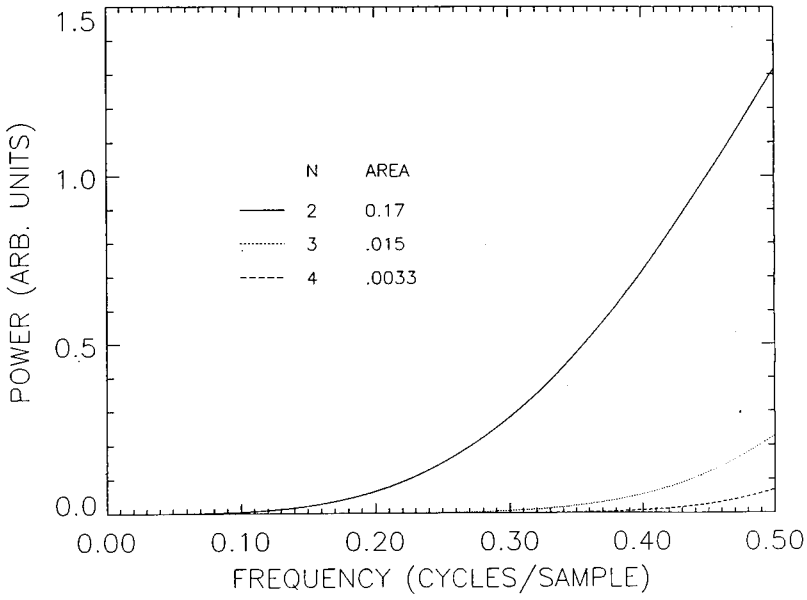

Fig. 10 — Error spectra for SLF-2, 3, and 4, and a $p = 2$ power law

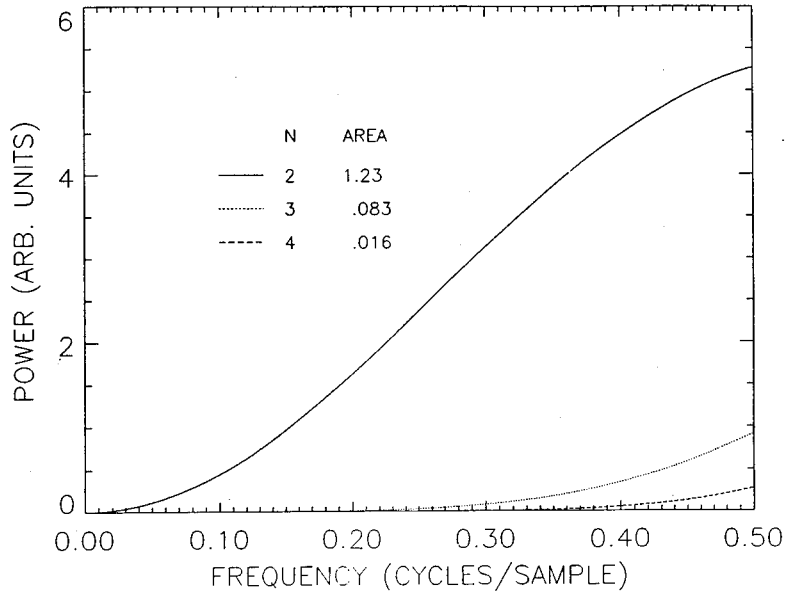

Fig. 11 — Error spectra for SLF-2, 3, and 4, and a $p = 4$ power law

Notice that the residual error spectrum plotted in Fig. 10 varies as ν^{4N-4} near dc, and as ν^{4N-6} in Fig. 11. Even for $N = 2$, this may be considered low-frequency overkill. The bulk of the error energy even for the $p = 4$ spectrum resides at high frequencies, which have been neglected in the SLF definition. The next subsection expands the filter design procedures to remedy this defect.

4.2 Hybrid Designs

The error factor in Eq. (53) and derivatives of it may be constrained to the value zero at other frequencies besides dc. For many of the DDFs and interpolators already discussed, the peak values of the error spectra occur at a high frequency, in particular, above the Nyquist. We now restrict the remaining discussion to oversampled systems, such as are assumed in Figs. 9 through 11. For them, the maximum error occurs at the Nyquist frequency, and a natural next step in filter design would be to sacrifice some low-frequency performance to reduce the residual error at higher frequencies.

The elimination of errors at nonzero frequencies is possible with interpolation as well as with DDFs and has been studied [12]. Notice, however, that $E_s(\nu)$ for interpolation (Eq. (44)) is complex at most frequencies, and requiring it to be zero usually expends two degrees of design freedom. An exception occurs at dc, at which frequency requiring Eq. (44) to be zero at all shifts s constitutes but one real constraint. On the other hand, because at $\nu = 1/2$ the transfer function $H_1^{(s)}$ of Eq. (42a) is real, $E_s(\nu)$ in Eq. (44) (i.e., for interpolation) can never be zero at all shifts. This limitation occurs because a sine wave at the Nyquist frequency is zero at all integer sample values, and is, therefore, invisible to the sampling process. Such a sinusoid cannot, therefore, be reproduced by any interpolator.

On the other hand, for symmetric DDFs, $E_s(\nu)$ may be taken as real (see Eq. (53)). Therefore, each frequency null requires the expenditure of only one degree of freedom. Furthermore, because of Eq. (52), the transfer functions in Eq. (42) for the two filters are complex conjugates, and the complex error factor in Eq. (43) may be written as

$$E_s(\nu) \rightarrow 2\left|H_1^{(s)}(\nu)\right| \sin\left(\phi^{(s)}(\nu) - \frac{\omega s}{2}\right), \tag{74}$$

in which $\phi^{(s)}$ is the phase of $H_1^{(s)}$. A unimodular factor has again been omitted.

Now note from Eq. (45) that if $H_1^{(s)}$ in Eq. (74) is derived from the sinc kernel at the shift $s/2$, then ϕ has the value $\omega s/2$ inband, making $E_s(\nu)$ zero inband for all shifts. This simply confirms that shifting a pair of discrete oversampled signals toward each other by half their separation gives the ideal result if the ideal interpolator can be used. However, Eq. (74) shows that zero inband error for DDFs does not require use of the ideal interpolator, but only one with the ideal phase. By contrast, for the interpolation problem, zero inband error in Eq. (44) requires specific use of the sinc interpolator.

According to Eq. (74), $E_s(\nu)$ can be made zero either by making the magnitude of the transfer function zero, or by adjusting its phase appropriately. Furthermore, this can be accomplished with DDFs even at the Nyquist frequency, an impossibility for interpolators.

As an example, we consider three-point DDFs for which the error is zero at angular frequencies ω_1 and ω_2. The form in Eqs. (53) and (56) can be written as

$$\begin{bmatrix} X_{-1}^1 & X_0^1 & X_1^1 \\ X_{-1}^2 & X_0^2 & X_1^2 \\ t_1 & 1 & t_{-1} \end{bmatrix} \begin{bmatrix} k(-1+s) \\ k(s) \\ k(1+s) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{75}$$

where

$$X_n^i = \sin\left[\omega_i\left(n + \frac{s}{2}\right)\right] \qquad (i = 1, 2;\ n = -1, 0, 1)\,. \tag{76}$$

According to Cramer's Rule, the solution $k(n + s)$ can be derived by substituting the right-hand vector in Eq. (75) for the nth column of the left-hand matrix, finding the determinant of this modified matrix, and dividing by the determinant of the original matrix. However, the zeroes in the vector make each such modified matrix independent of all t_n. Therefore, any solutions $\{k(n + s)\}$ for one set $\{t_n\}$ are proportional to those for any other, for example for $\{t_n = 1\}$. We may, therefore, solve Eq. (75) for $t_n = 1$ and rely on Eq. (70) for the physically important quantity, peak-signal/clutter. Note that because the SLF-3 filter may be viewed as the limiting solution to Eq. (75) as the frequencies $\nu_1, \nu_2 \to 0$, the prescription of Eq. (70) also applies to it and, more generally, to SLF-N.

The solution to Eq. (75) with $t_n = 1$ is

$$\begin{aligned} k(s-1) &= m_{31} \,/\, det \\ k(s) &= -m_{32} \,/\, det \\ k(s+1) &= m_{33} \,/\, det \end{aligned} \tag{77}$$

with m_{ij} the cofactors of the matrix in Eq. (75) and *det* its determinant.

Figure 12 plots the error factor for the above kernel with $\nu_1 = 1/4$, $\nu_2 = 1/2$. We will generally use the notation $SN(\nu_1, \nu_2, ...)$ to mean a symmetric N-point DDF with zeroes in its error factor at frequencies $\nu_1, \nu_2, \ldots$. If the number of arguments of SN is less than $N - 1$ — the single dc constraint of Eq. (69) is always assumed — then the remaining degrees of freedom are understood to have been used to set odd low-order derivatives of $E_s(\nu)$ equal to zero at $\nu = 0$. (Even ones are automatically zero.) Thus, $S3(1/4)$, whose performance is also depicted in Fig. 12, satisfies $E_s(1/4) = 0$ as well as $E_s'(0) = 0$. Such hybrid filters generally have low-frequency properties similar to SLFs, along with null errors at selected higher frequencies. As with SLFs, for hybrid filters the dc constraint may be assumed and Eq. (70) used to find S/C ratios.

Figure 12 also shows the performance of an $N = 4$ hybrid interpolator studied in Ref. 12. In this case, two of the interpolator's four degrees of freedom are used to create the null at $\nu = 1/4$; the other two make $E_s(0) = E_s'(0) = 0$. Therefore, at low ν, $|E_s|$ varies as ν^2 for the interpolator. For $S3(1/4, 1/2)$, $|E_s|$ varies as ν, performing worse than the interpolator at low frequencies, but better at all frequencies beyond $\nu \approx 0.13$. The error factor for $S3(1/4)$ begins with a ν^3 dependence and remains superior to the interpolator at all frequencies, but becomes worse than the other DDF above $\nu \approx 0.39$.

Figure 13 plots the corresponding error spectra for the $p = 2$ power spectrum. Both DDFs give smaller residual rms differences than the interpolator ($S3(1/4, 1/2)$ by the factor 3.4 and $S3(1/4)$ by the factor 4.5).

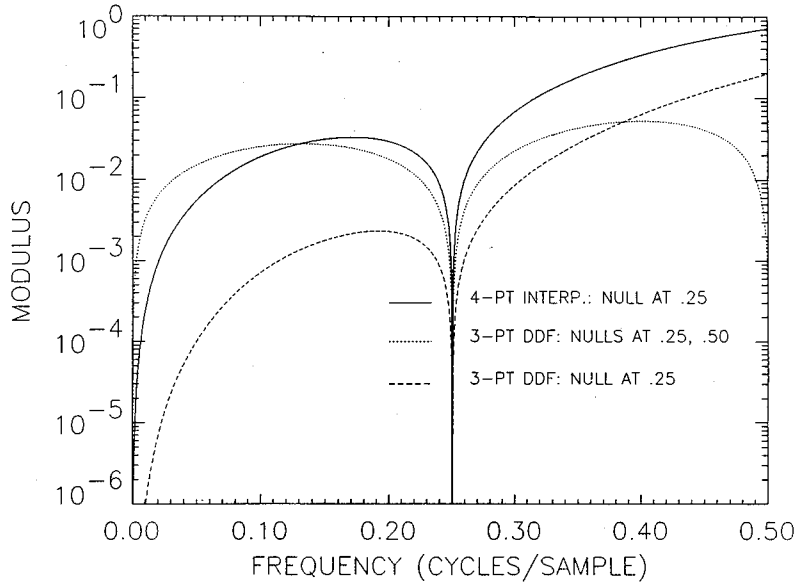

Fig. 12 — Errors factors for three hybrid methods. Besides at the frequency $\nu = 0$, other nulls occur at 1/4 for the interpolator and for both DDFs, $S3(1/4, 1/2)$ and $S2(1/4)$; also at $\nu = 1/2$ for $S3(1/4, 1/2)$.

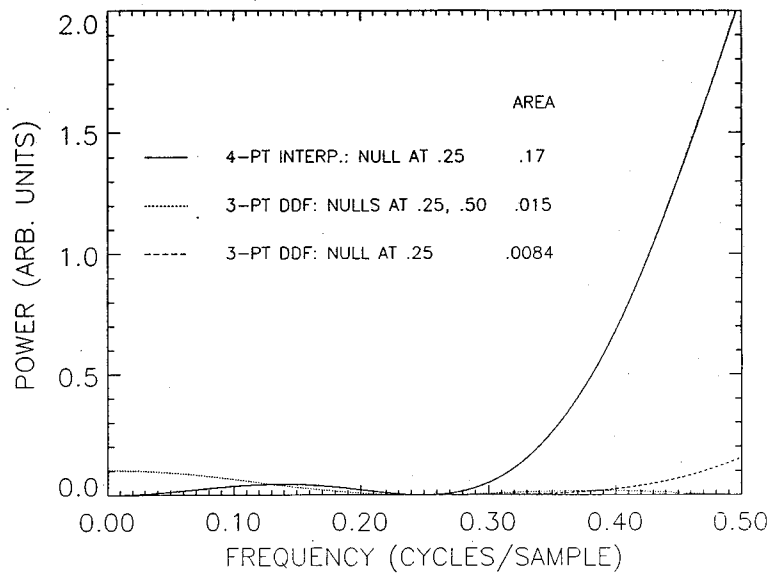

Fig. 13 — Error spectra for three hybrid methods and a $p = 2$ spectrum. The interpolator is of order 4; both DDFs are of order 3.

The relative gain in performance is even more impressive if the order of the DDF is as high as that of the interpolator. Figure 14 compares the error factors for $S4(1/4, 1/2)$ with the two DDFs of Fig. 12. The extra degree of freedom in the four-point method allows it to combine the low-frequency behavior of $S3(1/4)$ with the high-frequency behavior of $S3(1/4, 1/2)$. The corresponding error spectra are shown in Fig. 15. Figure 16 summarizes the $p = 2$ results and includes results for $S3(0.15, 0.4)$ and $S4(0.1, 0.3, 0.45)$, which have been found to be good general purpose DDFs for imaging applications. Similar gains are found for $p = 3$ and $p = 4$ power-law spectra.

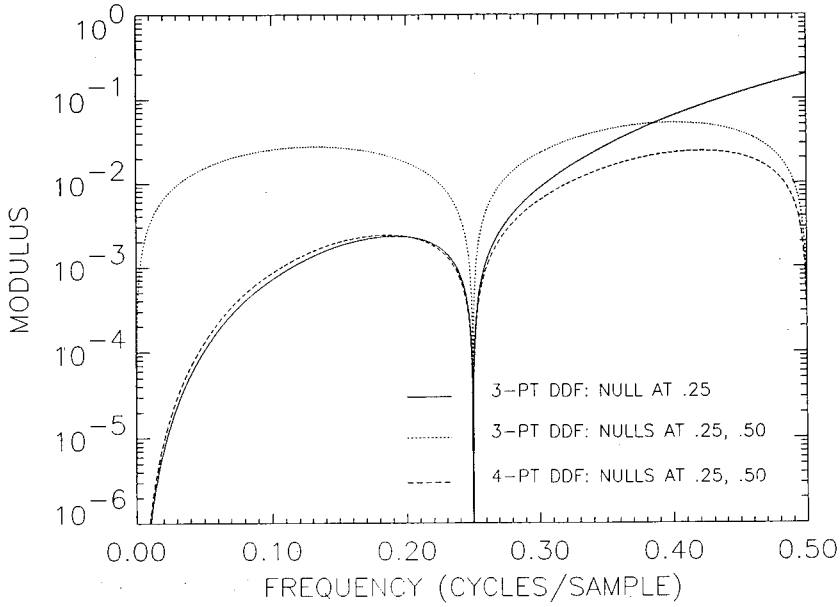

Fig. 14 — Comparison of hybrid DDFs

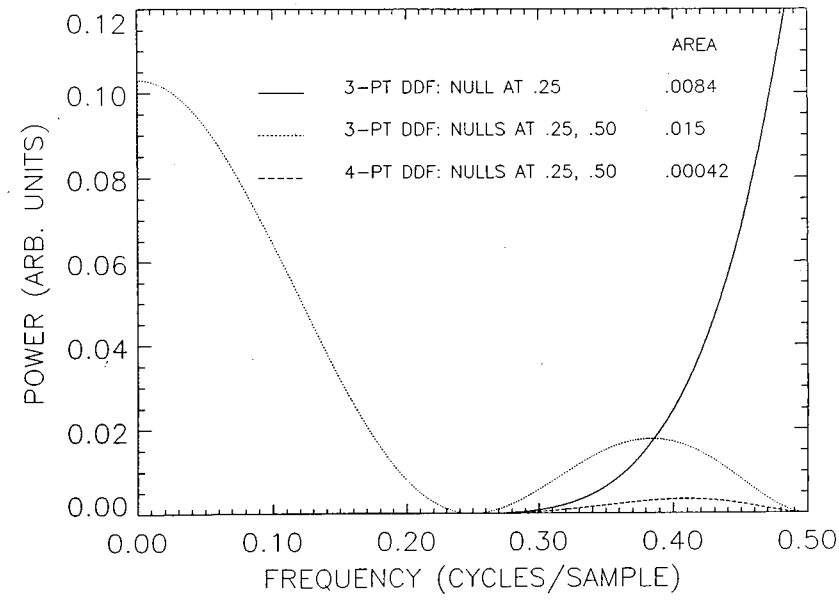

Fig. 15 — Error spectra for hybrid DDFs and a $p = 2$ input spectrum

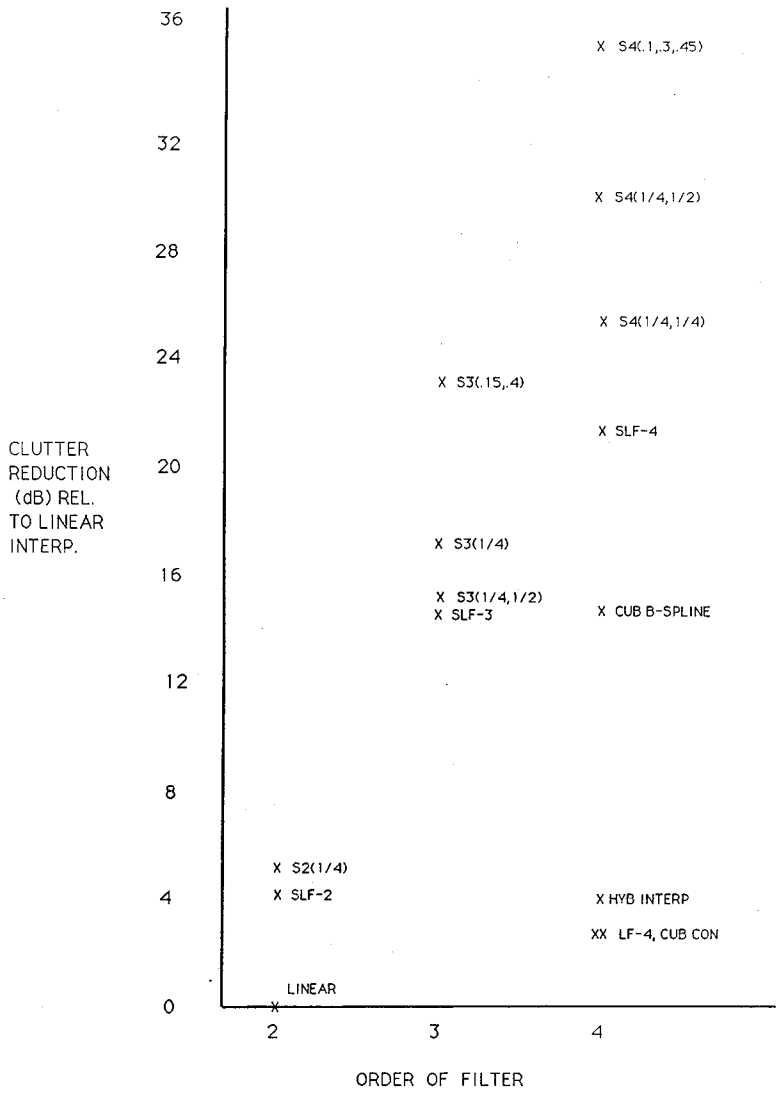

Fig. 16 — Performance of interpolators and DDFs relative to Linear, for a $p = 2$ input spectrum truncated at the Nyquist frequency. Numbers in parentheses indicate locations of null frequencies.

4.3 Effects on Gaussian Targets

Figures 17 through 20 illustrate the effects of two four-point DDFs on Gaussian targets. Figures 17 and 18 use the $p = 2$ optimal DDF, with Gaussian target constraints in Eqs. (56) and (57) for $\sigma = 0.5$ and 1.0, respectively. Figures 19 and 20 show similar results for the DDF $S4(1/4,1/2)$. The targets are assumed to have moved a distance of at least a few samples between the collection times of the discrete frames of data, so that they are not destroyed by the frame subtraction operation. The solid lines represent the Gaussians, which are sampled at points on the abscissa separated by a distance of one. If, for example, a Gaussian is sampled at its peak, located at position 5 in the plots, then the filtered target has a peak value identical to the unfiltered, because the two curves intersect there, a consequence of imposing the condition of Eq. (56) on the filters. If, on the other hand, a Gaussian signal happens to be sampled off its peak, then the corresponding value produced by the filter is usually higher than the input. That is, the filtered curves are usually above the unfiltered at any given sampling position.

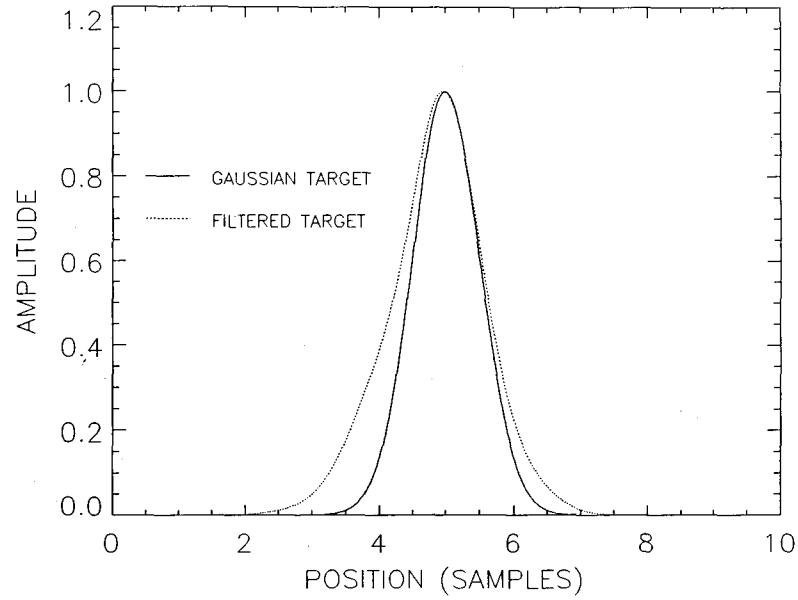

Fig. 17 — Effect on a $\sigma = 0.5$ Gaussian target of the optimal four-point DDF, designed with a $\sigma = 0.5$ Gaussian constraint, for a $p = 2$ power-law input spectrum and the shift $s = 0.25$ samples

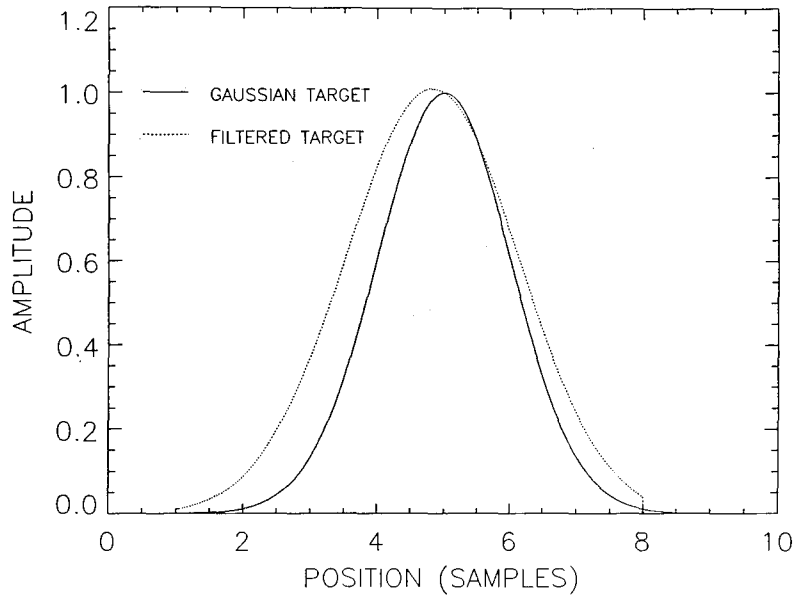

Fig. 18 — Effect on a $\sigma = 1.0$ Gaussian target of the optimal four-point DDF, designed with a $\sigma = 1.0$ Gaussian constraint, for a $p = 2$ power-law input spectrum and the shift $s = 0.25$ samples

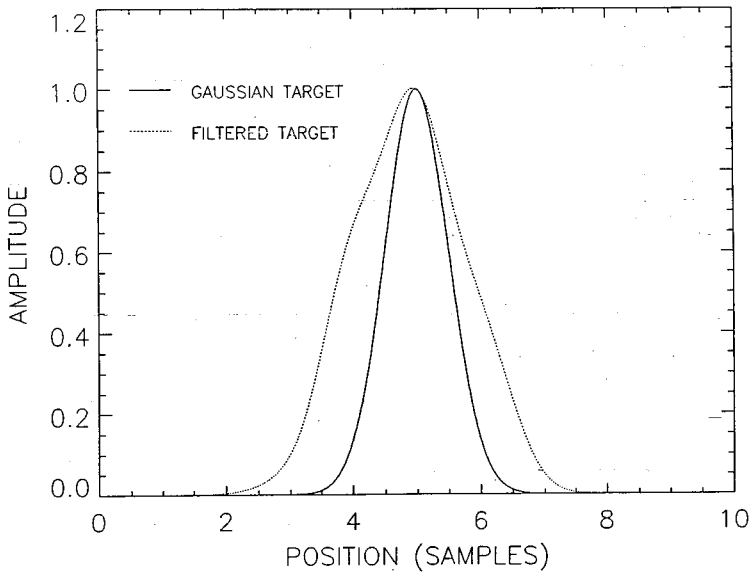

Fig. 19 — Effect of $S4(1/4, 1/2)$ DDF with $\sigma = 0.5$ Gaussian constraint on a $\sigma = 0.5$ Gaussian target

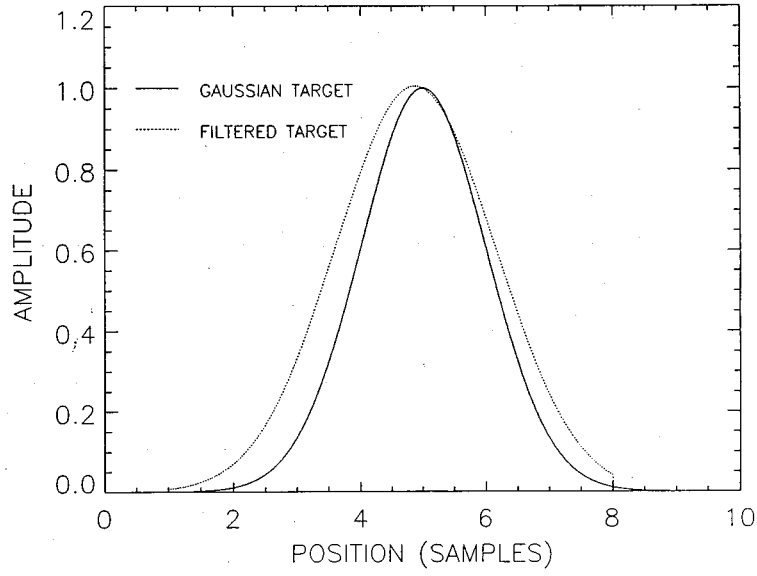

Fig. 20 — Effect of $S4(1/4, 1/2)$ DDF with $\sigma = 1.0$ Gaussian constraint on a $\sigma = 1.0$ Gaussian target

Thus, the signal-to-clutter formula (Eq. (70)) is usually a slight underestimate of the true gain. For example, if in Fig. 17 the sampling points are at the half integers, then the peak input signal is $\exp(-(1/2)^2/(2\sigma^2)) = 0.607$, whereas the peak filtered value is ≈ 0.75. Therefore, Eq. (70) actually underestimates the peak-signal/clutter by the factor 1.23. Note also that a second, filtered version of a moving Gaussian target appears in the difference frame; for symmetric DDFs, it corresponds to the negative mirror image of the appropriate dotted curve in Figs. 17 through 20.

More generally, the performance of the single-frame filters can be characterized by the usual frequency descriptor of digital signal processing, the magnitude of the transfer function $H^{(s)}$. Figure 21 plots these for three three-point DDFs with the dc constraint. These show the typical low-pass characteristic of all the filters discussed here.

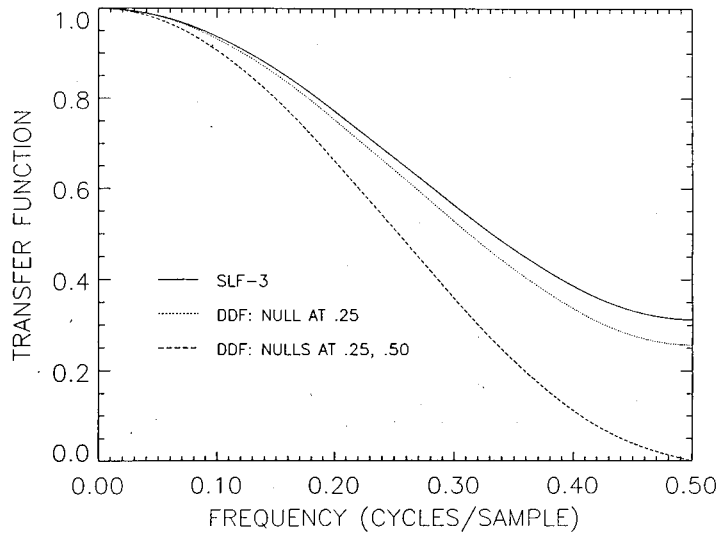

Fig. 21 — Transfer functions of three three-point DDFs with dc constraint

Figure 22 shows, for each of the three constraints, the transfer functions for an effective inband DDF, the hybrid $S4(1/4,1/2)$. Note that four of the six filters described in Figs. 21 and 22 have difference-error nulls at $\nu = 1/2$, five have them at $\nu = 1/4$, and all have them at $\nu = 0$. However, these nulls are achieved differently, according to Eq. (74). The null error factors at $\nu = 1/2$ are always zero because $|H^{(s)}| = 0$. By contrast, the plots show that $H^{(s)}$ is never zero at the two lower frequencies, and so these nulls are achieved by having $\phi^{(s)} = \omega s/2$.

Note also that the three transfer functions in Fig. 22 are proportional, a reflection of the previously described proportionality of the corresponding kernels $\{k(n+s)\}$ for hybrid DDFs. This should be contrasted with the optimal four-point filter for the $p = 2$ spectrum in Fig. 23, where the dependence on the inhomogeneous target constraint of Eq. (56) is more complicated. Finally, in each of these two figures, the two dc amplifications occur because of the combination of the low-pass filter characteristics of $H^{(s)}$ and the constraint preserving peak value of a Gaussian target. These require the filter to broaden the target without depressing its central value. The dotted curves in Figs. 17 through 20 all have larger means than the Gaussians that generated them.

5. RELATED ISSUES

Reference 16 discusses several corollaries of the mean squared-error formulation developed here. These include generalizations to higher dimensions, in which the variables x, n, and the shift s become vectors. Also, an extension of the analysis to include more than two signals is achieved which, for image processing applications, transforms the problem into one of multiframe tracking. Generalizations in Ref. 16 are also developed that allow the evaluation and optimization of methods of prediction, restoration, and finite-length matched filtering.

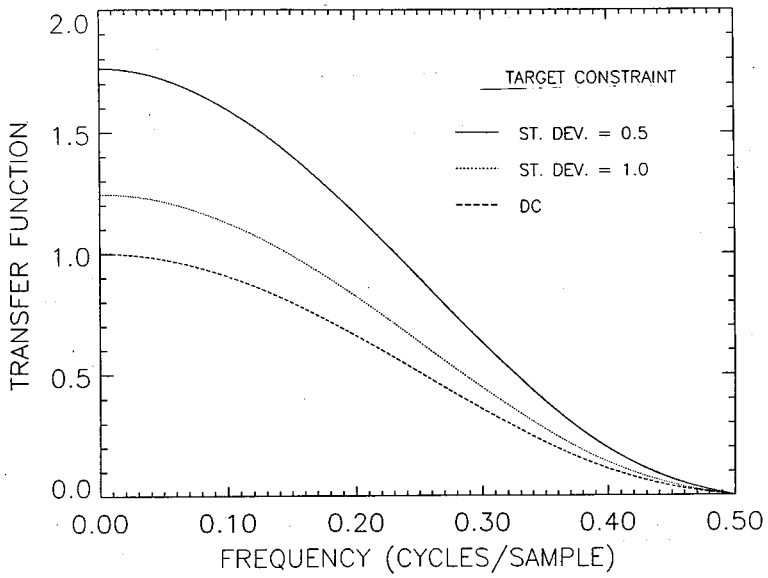

Fig. 22 — Transfer functions of *S*4(1/4, 1/2) for three different target constraints

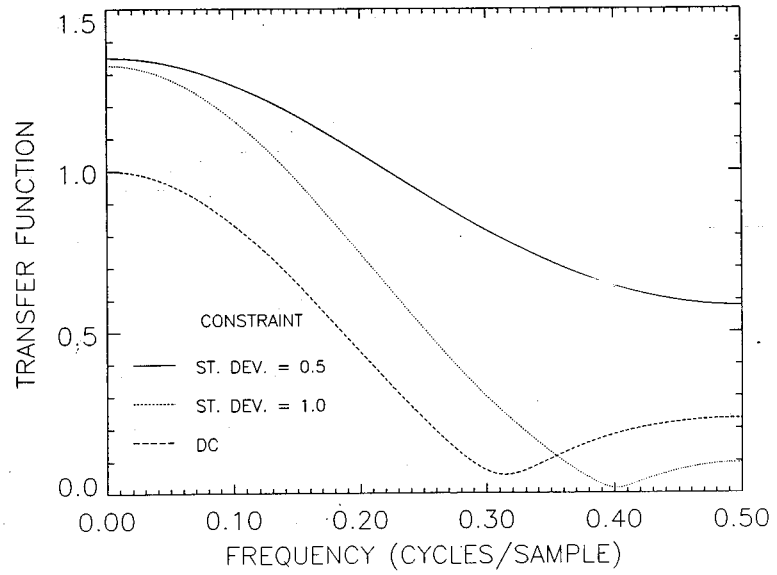

Fig. 23 — Magnitude of transfer functions of the four-point optimal (for a $p = 2$ input power-law spectrum) DDFs corresponding to three different target constraints

Here we include a discussion of additive temporal noise. This report is concerned mostly with the error d_s, defined in Eq. (25). In image processing applications, d_s is called residual clutter. Initial clutter is just the variance of the unprocessed image. According to Eqs. (41) and (36), if the second filter k_2 vanishes and $K_1^{(s)}(n)$ is δ_0^n, i.e., if nothing is done to the first image, then $|E_s|= 1$ and d_s^2 is the variance of the image (by Plancherel's formula [19]). Therefore, according to Eq. (36), the squared modulus of the error factor is the ratio of the power in the filtered difference spectrum to that in the original image.

In images for which stochastic (temporal) noise appearing in both images is comparable to the spatial clutter, or for processed difference images in which the residual clutter has been reduced to the noise level, Eq. (36) is inaccurate. Stochastic noise contributions $\mathcal{N}$ must be added to the digital signals, so that the replacements

$$f(n) \rightarrow f(n) + \mathcal{N}_1(n)$$

$$f(n+s) \rightarrow f(n+s) + \mathcal{N}_2(n) \tag{78}$$

are required.

Note that if the second image is unaltered, in which case d_s is a measure of interpolation accuracy, then the methodology of Sections 2 and 3 applies to the restoration of a digital image with additive noise $\mathcal{N}$. In this case, the problem becomes one of finite Wiener filtering, and the optimization equation describes the minimum mean-squared error finite digital filter. Note that this is usually not simply a truncated version of the standard Wiener solution, which involves an infinite-order filter $k(n+s)$.

Here we assume that the noise terms in Eq. (78) are independent both of each other and of the image f, and that they are zero-mean stationary ergodic random processes with the following common autocorrelation function $\mathcal{A}$:

$$\mathcal{A}(n) = \mathcal{E}\left\langle \mathcal{N}_i(m)\,\mathcal{N}_i(n-m)\right\rangle , \tag{79}$$

where $\mathcal{E}$ denotes an ensemble average for $i = 1$ and 2. Then Eq. (54) generalizes to

$$d_s^2 = 2\sum_{m,n}\left\{\left[R(n)+\mathcal{A}(n)\right]k(n+m+s)\,k(m+s) - R(n+s)\,k(n+m+s)\,k(-m+s)\right\} , \tag{80}$$

which includes both clutter and noise. The optimization (Eq. (55a)) is replaced by

$$\sum_m \left[R(m)+\mathcal{A}(m)\right]k(n-m+s) - \sum_m R(m+s)\,k(m-n+s) = \sum_i \lambda_i \frac{\partial c_i}{\partial k(n+s)} . \tag{81}$$

As an example of differences between optimal solutions with and without noise, we consider $N = 2$ solutions for the Lorentzian clutter spectrum, with the general target constraint of Eq. (56) (with $t_0 = 1$). The temporal noise is assumed to be white, i.e., $\mathcal{A}(m) \propto \delta_0^m$, and r is the noise to clutter ratio $r = \mathcal{A}(0)/R(0)$. The optimal DDF is given by

$$k(-1) = \frac{\rho - \rho^{1-s} + t_1(\rho^s - 1 - r)}{\rho^{2-s} - 1 - r - 2t_1(\rho - \rho^{1-s}) - t_1^2(1 - \rho^s + r)}$$

$$k(0) = 1 - t_1 k(-1) , \tag{82}$$

with ρ given by Eq. (61).

If the noise were absent, then $r = 0$ and the solution as $s \to 0$ becomes

$$k(-1) = 0$$

$$k(0) = 1 . \tag{83}$$

This corresponds to complete removal of clutter by the DDF by reproducing the sample values of both digital signals. If the solution in Eq. (82) with $r = 0$ is used to reduce a severe clutter problem that might exist at large shifts — that is, if noise is ignored — then at small shifts it can be noticeably suboptimal, as measured by the ratio of signal to clutter-plus-noise. In particular, if the noise term is kept in Eq. (82), the solution at $s = 0$ is

$$k(-1) = \frac{r t_1}{1 - \rho^2 + r(1 + t_1^2)}$$

$$k(0) = 1 - t_1 k(-1) . \tag{84}$$

Noise cannot be totally removed at any shift, and this solution represents a balance in which the filter keeps some clutter and some noise. If noise happens to dominate clutter, i.e., r is large, then Eq. (84) approaches

$$k(-1) = \frac{t_1}{1 + t_1^2}$$

$$k(0) = \frac{1}{1 + t_1^2} , \tag{85}$$

which is the two-point matched filter in white noise. This DDF produces a smaller (mean squared) signal-to-noise ratio than does the DDF of Eq. (83) by the factor $(1 + t_1^2)$. Such is the extent to which the DDF designed purely for clutter suppression is suboptimal, when it is applied in a noise-dominant environment.

6. SUMMARY AND CONCLUSIONS

This report generalizes the notion of interpolation and digital subtraction as a means of detecting changes in a pair of discrete signals. By acting on both signals, Dual Difference Filters afford more flexibility and power than do interpolators of the same order. This is

proven by a fundamental theorem (Eq. (36)) of mean squared error for the general resampling problem, which includes interpolation as a special case.

For any fixed shift between sampling grids, the strong form of the theorem associates an error spectrum with any DDF that is applied to an oversampled signal. The error spectrum is the product of the power spectrum of the continuous underlying signal that generates the discrete samples, and an attenuation factor. The weak form of the theorem says the same thing for undersampled signals, but in an average sense. Formulae are derived for the attenuation factors that can be used for filters that are of finite extent in either frequency-space (Eq. (37)) or real-space (Eq. (41)). A corollary of the weak theorem allows application of the error formulae to stochastic problems.

A fundamental optimization equation (Eq. (58)) is derived. If the precise form of the signal spectrum is known, then the solution of the equation produces the maximum possible peak-signal to rms clutter for any DDF of given order. As an example, the case of a power-law spectrum and a single linear constraint on the filters is solved.

Flexible design methods for DDFs are developed that produce superior performance for a wide range of signal spectra. For example, for three bandlimited power-law spectra ($p = 2$, 3, and 4), a single four-point DDF is found that enhances residual signal/clutter by approximately 30 dB, compared to the best popular interpolators of the same order.

REFERENCES

1. G. W. Wornell, D. S. Kauffman, and B. Sharpe, "Minimum Mean-Square Error Resampling for Remotely Sensed Imagery," *12th Canadian Symposium on Remote Sensing, IEEE* **89CH2768-0**, 1989.

2. J. Owczarczyk, W. J. Welsh, and S. Searby, "Performance Analysis of Image Registration Techniques," *Third Int. Conf. on Im. Proc. and Applic.*, July 18-20, 1989, Electronics Div. of the Institution of Electrical Engineers, University of Warwick, UK, pp. 10-13.

3. J. A. Parker, R. V. Kenyon, and D.E. Troxel, "Comparison of Interpolating Methods for Image Resampling," *IEEE Trans Med. Im.* **MI-2**, 31-39 1983.

4. K. P. Prasad and P. Satyanrayana, "Fast Interpolation Algorithm Using the FFT," *Electron. Lett.* **22**, 185-187, 1986.

5. A. J. Mord, N. H. Endsley, E. Ramberg, and H.J. Reitsema, "Electronic Image Alignment: Implementation and Applications to Imaging System Design," *Opt. Eng.* **24**(3), 507-515, 1985.

6. A. Ferraioli, "Review of Clinical Applications of Cardiac Digital Angiography," *IEEE Trans. Med. Im.* **9**(1), 113, 1990.

7. A. Schaum, "Dual Difference Filtering for MTI in Digital Imagery," IRIS Specialty Group on Targets, Backgrounds, and Discrimination, Jan. 30, 1992, Naval Training Center, Orlando, FL.

8. A. Schaum and M. McHugh, "Analytic Methods of Image Registration: Displacement Estimation and Resampling," NRL Report 9298, Feb. 1991.

9. A. Stocker and A. Oshagan, "Comparison of Interpolation Kernels," Report SCC-R-181-5, Space Computer Corp., Santa Monica, CA, Feb. 1991.

10. A. Schaum, "Principles of Interpolator Design and Evaluation," NRL Report 9356, Nov. 1991.

11. R. N. Bracewell, *The Fourier Transform and its Applications* (McGraw-Hill, New York, 1978).

12. A. Schaum, Theory and Design of Local Interpolators, submitted to *Comput. Vision, Graphics, and Im. Proc.*

13. G. Oetken, T. Parks, and H. Schüssler, "New Results in the Design of Digital Interpolators," *IEEE Trans. Acoust. Sp. Sig. Proc.* **ASSP-23**(3), 1975.

14. E. A. Robinson and S. Treitel, *Geophysical Signal Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1980), p. 163.

15. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes* (Cambridge Univ. Press, Cambridge, 1986), pp. 47-51.

16. A. Schaum, "Dual Difference Filtering for Change Detection in Digital Signals," for publication in *IEEE J. Sig. Proc.*

17. S.M. Kay, *Modern Spectral Estimation: Theory and Application* (Prentice-Hall, Englewood Cliffs, NJ, 1988).

18. S. K. Park and R. A. Schowengerdt, "Image Reconstruction by Parametric Cubic Convolution," *Comput. Vision, Graphics, and Im. Proc.* **23**, 1982.

19. R. Seeley, *An Introduction to Fourier Series and Integrals* (W.A. Benjamin Inc., New York, 1966).